

Fragments of nucleic acids specific to mycobacteria  
which are members of the M. tuberculosis complex and  
their applications for the detection and the  
differential diagnosis of members of the M.  
5 tuberculosis complex

The present invention relates to sequences of nucleic acids of mycobacteria belonging to the M. tuberculosis complex.

10 The invention likewise relates to a method of in vitro detection of strains of mycobacteria belonging to the M. tuberculosis complex as well as to a method for a differential diagnosis of strains of the M. tuberculosis complex, especially for differentiating  
15 the presence of the BCG from that of other members of the complex in a sample.

Approximately, 1.7 billion people or 1/3 of the world population are infected with M. tuberculosis (Sudre et al., 1992). In 1990, the estimated number of  
20 cases of tuberculosis was 8 million, including 2.9 million deaths (Sudre et al., 1992). These last few years, the number of cases of tuberculosis in the United States and in Europe has increased by 3 to 6% per annum, principally in populations at high risk such  
25 as patients suffering from AIDS, chronic alcoholics, the homeless and drug addicts (Barnes et al., 1991).

Taking account of the difficulties in fighting infections by mycobacteria, there is an urgent need to be able to have a specific and sensitive rapid method  
30 allowing these infections to be diagnosed. Also, the early detection of M. tuberculosis in clinical samples is taking on growing importance in the control of tuberculosis both for the clinical treatment of infected patients and for the identification of exposed  
35 individuals at risk.

Detection by PCR (Polymerase Chain Reaction) of specific DNA of species of mycobacteria is probably one of the most promising novel approaches for rapid, specific and sensitive diagnosis (Saiki et al., 1985:

10086206.02802

- 2 -

10086206-022802

5   Brisson-Noël et al., 1989; Cousins et al., 1992; Eisebach et al., 1990; Forbes et al. 1993; Fries et al., 1991; Hermans et al., 1990; Jonas et al., 1993; Kolk et al., 1992; Pierre et al., 1991; Saboor et al., 1992; Shankar et al., 1991; Sjöbring et al., 1990). The different studies carried out up to date, however, have led to different results as far as the specificity and the sensitivity are concerned. Among the reasons for this diversity, it is especially possible to note  
10   methodological differences concerning the preparation of the samples, the protocol followed for the amplification or the methods of detection of the PCR products.

15       Several of these studies relate to the detection by PCR of the *M. tuberculosis* complex starting from the target DNA IS6110 (Clarridge et al., 1993; Eisenach et al., 1990 ; Forbes et al., 1993; Noordhoek et al., 1994), the gene coding for the 65 kDa antigen (Brisson-Noël et al., 1991; Telenti et al.,  
20   1993), the gene coding for the 38 kDa antigen (Folgueira et al., 1993; Sjöbring et al., 1990) or the ribosomal 16S RNA (Kox et al., 1995). However, all these diagnostic tests identify the *M. tuberculosis* complex in its entirety.

25       The *M. tuberculosis* complex comprises *M. tuberculosis*, *M. bovis*, *M. microti* and *M. africanum*. These four species have strong homologies in their DNA (85 to 100%) (Imeada, 1985) and have several totally identical genes. This homology has limited the use of  
30   DNA sequences to differentiate the strains. It would nevertheless be of particular interest to be able to differentiate the strains of *M. tuberculosis* from those of the Calmette-Guérin (BCG) bacillus *M. bovis*, the latter often being used as live vaccines for  
35   immunoprotection against tuberculosis. Obviously, there is thus an interest in distinguishing between possibly pathogenic mycobacteria and vaccinating BCG strains. This distinction is particularly important in the case

- 3 -

of immunodeficient individuals, such as subjects infected by HIV.

Restriction fragment length polymorphism (RFLP) analyses based on the insertion of the IS6110 elements have been used to differentiate different strains of *M. tuberculosis*. It has been shown that this insertion element allowed specific RFLP profiles of strains to be obtained on account of a variability in its localization and in the number of copies existing in the genomes of different strains. IS6110 has been demonstrated in *M. tuberculosis* and in *M. bovis* but not in the other mycobacteria tested (Cave et al., 1991). In general, several copies of this insertion element can be demonstrated in *M. tuberculosis*, although a single copy is found in *M. bovis* BCG (Cave et al., 1991). However, certain strains of *M. tuberculosis* are devoid of IS6110 (van Soolingen et al., 1994) and strains of *M. bovis* having a significant number of copies are common in some populations (van Soolingen et al., 1994). Added to these limitations is the fact that the identification of strains of *M. tuberculosis* from IS6110 necessitates Southern Blot analyses and cannot easily be carried out by routine PCR methods.

The authors of the present invention have demonstrated a novel target DNA sequence for enzymatic amplification. This sequence is part of an operon coding for a specific two-component regulatory system of *M. leprae* and bacteria which are members of the *M. tuberculosis* complex. The two-component systems are regulatory systems belonging to a large family, involving two proteins which cooperate in translating external signals by modifications in the level of genetic expression (Parkinson and Kofoed, 1992). According to a proposed general model, one component localized in the membrane could act as a sensor of environmental changes and could transmit the information to a regulatory component of the response in the cytoplasm, which in turn could modulate the

20086206 DEB02

- 4 -

transcription of target genes. The communication between the two components is generally carried out by a cascade of phosphorylations.

The authors of the present invention have at present cloned and characterized a two-component mycobacterial system. This system appears to be specific to members of the *M. tuberculosis* complex and to *M. leprae*. Unexpectedly with respect to the other two-component systems, the genes of this mycobacterial system are separated by DNA sequences (intercistronic sequences or repeated sequences) which are uniquely present among the strains of mycobacteria of the *M. tuberculosis* complex and in *M. leprae*.

The authors of the present invention have additionally shown that in the strains of the *M. tuberculosis* complex, this intercistronic region corresponded to an exact or truncated number of repetitions of a sequence of 77 base pairs, SEQ ID No. 1, which was able to vary among the strains. The truncated sequence (SEQ ID No. 2) is composed of 53 base pairs and contains a short in-phase internal deletion of nucleotides Nos. 40 to 66, which are substituted by a GAG codon.

The authors of the invention have likewise shown that, surprisingly, the truncated sequence (SEQ ID No. 2) was characteristic of the members of the *M. tuberculosis* complex which are different from BCG.

In *M. leprae*, the corresponding intercistronic region is formed by a variant of 52 base pairs already additionally described and whose sequence is indicated below:

5' atg aca ccc gcg cag gcg atg atg cag agc gaa gtg acg aga ggg aat gtg a 3'.

Taking account of their characteristics, these repeated sequences are of particular interest for identifying and differentiating between strains of the *M. tuberculosis* complex by enzymatic amplification techniques, such as PCR or other similar methods. More

10086206-022802

- 5 -

particularly, they allow a differential diagnosis between the presence of BCG\* and that of other members of the complex in a biological sample to be established. This differential diagnosis, whose principle is based on the specific detection of the sequence SEQ ID No. 2, forms a preferred aspect of the invention. It is advantageously employed to distinguish an infection by BCG from an infection by other virulent members of the complex in immunodeficient individuals, such as especially subjects infected by HIV.

The invention thus relates to a specific fragment of nucleic acids of mycobacteria belonging to the *M. tuberculosis* complex, comprising a sequence of nucleotides chosen from amongst the sequence SEQ ID No. 1, the sequence SEQ ID No. 2, their complementary sequences or the sequences of nucleic acids capable of hybridizing with one of the preceding sequences under conditions of high stringency.

It is likewise aimed at the use of the said sequences for the production of nucleotide probes for the in vitro detection of mycobacteria belonging to the *M. tuberculosis* complex and of oligonucleotide primers for the enzymatic amplification of specific sequences of strains which are members of the *M. tuberculosis* complex.

The invention likewise relates to a method allowing the strains of the *M. tuberculosis* complex to be differentiated from one another, particularly for allowing BCG to be differentiated from pathogenic mycobacteria of this complex.

It likewise provides means for differentiating sub-groups among the strains forming the *M. tuberculosis* complex.

In addition, it can allow *M. leprae* of the species and of the strains belonging to the *M. tuberculosis* complex to be differentiated in a biological sample.

209220 90228007

- 6 -

In the context of the present invention, it is considered that the conditions of "high stringency" in which two nucleotide sequences can hybridize are the conditions defined by Sambrook et al., 1989, namely  
5 temperature conditions of between ( $T_m$  minus  $5^\circ\text{C}$ ) and ( $T_m$  minus  $30^\circ\text{C}$ ) and additionally preferably temperature conditions of between ( $T_m$  minus  $5^\circ\text{C}$ ) and ( $T_m$  minus  $10^\circ\text{C}$ ) (high stringency),  $T_m$  being the melting temperature, defined as being the temperature at which 50% of the  
10 paired strands separate. Preferentially, the conditions of high stringency used correspond to prehybridization, hybridization and washing temperatures of  $68^\circ\text{C}$  and a prehybridization and hybridization buffer based on 5 x SSC (protocol recommended by Boehringer Mannheim).

15 The invention relates to a specific fragment of nucleic acids of mycobacteria of the *M. tuberculosis* complex localized in the intercistronic region of the *senX3-regX3* system.

As will be explained in the examples, the  
20 sequence SEQ ID No. 1 corresponds to an entity of 77 repeated base pairs and a different number of copies according to the strains of mycobacteria studied. This repeated sequence has an open reading frame (ORF) which can code for a peptide of 25 amino acids.

25 The invention is thus aimed at a fragment of nucleic acids specific to the *M. tuberculosis* complex, comprising a sequence of nucleotides selected from the sequence SEQ ID No. 1, its complementary sequence or the sequences of nucleic acids capable of hybridizing  
30 with one of the preceding sequences under conditions of high stringency.

According to another aspect, the invention is aimed at a specific fragment of nucleic acids of members of the *M. tuberculosis* complex which are  
35 different from BCG, especially the virulent species *M. tuberculosis*, *M. africanum* or *M. bovis*, comprising a sequence of nucleotides chosen from amongst the sequence SEQ ID No. 2, its complementary sequence or

10086206-022802

- 7 -

the sequences of nucleic acids capable of hybridizing with one of the preceding sequences under conditions of high stringency.

As is actually illustrated by the examples, the BCG strains can be differentiated from all the other strains of the *M. tuberculosis* complex because of the absence of the sequence SEQ ID No. 1 in the *senX3-regX3* intergenic region in BCG (see Table 3, groups 4, 6, 8).

The different nucleotide sequences of the invention can be of artificial origin or non-artificial origin. They can be DNA or RNA sequences.

They can be prepared, for example, by chemical synthesis, or alternatively by mixed methods including chemical or enzymatic modifications of sequences obtained by screening sequence libraries, by means of probes elaborated on the basis of the sequences SEQ ID No. 1 or 2. Such libraries can be prepared by the classical techniques of molecular biology, known to the person skilled in the art.

The nucleotide sequences of the invention allow the production of probes or of nucleotide primers capable of hybridizing specifically with these, their corresponding RNA sequences or the corresponding genes under conditions of high stringency. Such probes are likewise part of the invention. They can be used as an in vitro diagnostic tool for the detection, by hybridization experiments, of specific nucleic acid sequences of mycobacteria belonging to the *M. tuberculosis* complex.

Preferentially, the probes of the invention have at least 24 consecutive nucleotides although shorter probes can be equally suitable. At the maximum, these have the complete *senX3-regX3* intergenic region of *M. tuberculosis* (IPL), namely two successive sequences SEQ ID No. 1 followed by a sequence SEQ ID No. 2. This DNA fragment contains 218 base pairs.

Among the specific preferred nucleotide probes of the members of the *M. tuberculosis* complex, the

10086306.022802

- 8 -

sequence SEQ ID No. 1 in its entirety or its complementary sequence appear especially.

According to another aspect, the invention is aimed at nucleotide probes for the detection and the demonstration of sequence of nucleic acids specific to members of the *M. tuberculosis* complex which are different from BCG, especially the virulent species *M. tuberculosis*, *M. africanum* or *M. bovis*, as well as for the differential diagnosis of the presence of BCG in a biological sample containing mycobacteria of the *M. tuberculosis* complex.

Such nucleotide probes are obtained from a region of the sequence SEQ ID No. 2 surrounding the GAG codon in position 40-42 or of its complementary strand. Preferentially, this region has a length of 21 base pairs, and contains 9 nucleotides upstream and downstream of the specific GAG codon of the sequence SEQ ID No. 2.

Advantageously, these probes comprise the sequence SEQ ID No. 2 in its entirety or its complementary sequence.

The probes of the invention hybridize according to the appropriate hybridization conditions, corresponding to the conditions of temperature and of ionic strength usually used by the person skilled in the art.

Preferentially, the probes of the invention are labelled, previously to their use. For this, several techniques are at the disposal of the person skilled in the art (fluorescent, radioactive, chemiluminescent, enzymatic, etc. labelling).

According to a preferred embodiment, the probes are labelled with digoxigenin. Digoxigenin (DIG) is a steroid hapten coupled to dUTP for labelling DNA used as a probe. This technology is marketed by Boshringer Mannheim. After the hybridization with the probe and washing, the detection is carried out following an emission of a chemoluminescent signal by means of the

10086206-022802



- 9 -

substrate disodium 3-(4-methoxyspiro(1,2 dioxethane-3,2'-(5'-chloro)tricyclodecane)-4-yl)phenylphosphate (CSPD).

The nucleotide sequences of the invention are likewise useful for the production and the use of oligonucleotide primers for sequencing reactions or for enzymatic amplification.

The enzymatic amplification techniques are principally illustrated by PCR. Other similar methods can, however, be used such as, for example, LCR (Ligase Chain Reaction), NASBA (Nucleic Acid Sequence Based Amplification), Q- $\beta$ replisase, SDA (Strand Displacement Amplification) and any other variant comprised in the technological knowledge of the person skilled in the art. These nucleic acid amplification techniques use oligonucleotide primer molecules to initiate the elongation reaction of the target sequence. The exact length of these primers will be able to vary according to the case. For example, as a function of the complexity of the sequence of the matrix, a polynucleotide primer typically contains from 15 to 25 nucleotides or more. In certain cases, however, it can contain less.

Preferentially, the nucleotide primers of the invention comprise at least 19 nucleotides. They are formed by primers chosen from sequences adjacent to the *senX3-regX3* intergenic region, in the regions 3' of *senX3* and 5' of *regX3*.

According to a preferred variant, the pair of primers referred to as C5 and C3 are used, namely 5'GCGCGAGAGCCCGAACTGC3' and 5'GCGCAGCAGAAACGTCAGC3' corresponding respectively to the 3' end of the *senX3* gene and the 5' end of the *regX3* gene. These primers respectively hybridize 56 base pairs upstream of the intercistronic region and 62 base pairs downstream of the latter.

The nucleotide sequences and the probes resulting therefrom can be cloned in cloning and/or

10086206.022802

- 10 -

expression vectors according to classical techniques of molecular biology, especially involving the use of restriction enzymes and of specific cleavage sites.

A preferred cloning vector in the present invention is represented by the plasmid pReqXMT1 deposited at the CNCM under the number I-1766 whose construction is described in greater detail in the examples below.

Another cloning vector according to the invention is represented by the plasmid pRegX3Bcl1 deposited at the CNCM under the number I-1765. These plasmids have each been introduced into the bacterium *E. coli* XL1-blue.

The vectors I-1765 and I-1766 each contain the complete *senX3-regX3* genes with the intercistronic regions of BCG and of *M. tuberculosis* respectively.

The nucleotide sequences according to the invention can additionally be expressed in appropriate systems, for the production and the study of the biological activities of the corresponding peptides. In this case, these will be placed under the control of signals allowing their expression in a cell host.

An efficient system of production of a protein or of a recombinant peptide necessitates having a vector, for example of plasmid or viral origin, and a compatible host cell.

The cell host can be chosen from prokaryotic systems, like bacteria, or eukaryotic systems, for example like yeasts, insect cells, CHO (Chinese hamster ovary) cells or any other system advantageously available.

The vector must contain a promoter, translation initiation and termination signals, as well as the appropriate regions of transcription regulation. It must be able to be maintained stably in the cell and can possibly have particular signals specifying the secretion of the translated protein.

- 11 -

These different control signals are chosen as a function of the cell host used. To this end, the nucleotide sequences according to the invention can be inserted into autonomic replication vectors in the chosen host, or integrative vectors of the chosen host. Such vectors will be prepared according to the methods currently used by the person skilled in the art, and the clones resulting therefrom can be introduced into an appropriate host by standard methods, such as, for example, electroporation.

The in vitro diagnostic or detection methods in which the nucleotide probes obtained from sequences of the invention are employed are likewise part of the present invention.

More particularly, the invention is aimed at a method of detection of strains of mycobacteria belonging to the *M. tuberculosis* complex in a biological sample comprising the following steps:

(i) contacting the biological sample with a pair of primers under conditions allowing hybridization of the said primers to the nucleic acids specific to strains of mycobacteria belonging to the *M. tuberculosis* complex;

(ii) amplification of the said nucleic acids;

(iii) contacting a nucleotide probe according to the invention with the said biological sample under conditions allowing the formation of hybridization complexes between the said probe and the amplified nucleic acid sequences;

(iv) detection of the hybridization complexes formed.

According to a first variant, the method according to the invention allows the presence of any member of the *M. tuberculosis* complex to be detected in a biological sample. In this case, the complexes of step (iii) defined above are formed with a specific nucleotide probe of the sequence SEQ ID No. 1 or of its complementary strand.

20086206-02202

- 12 -

According to a second advantageous variant, the method of the invention allows the presence of members of the *M. tuberculosis* complex other than BCG to be specifically detected. According to this variant, the complexes of step (iii) which are detected are formed with a nucleotide probe specific to the sequence SEQ ID No. 2 or to its complementary strand, consisting preferentially of a short sequence composed of two times 9 base pairs framing the GAC codon at the specific positions 40 to 42 of the sequence SEQ ID No. 2.

According to a third particularly advantageous variant, the method of the invention allows differential diagnosis of the BCG and of other members of the complex. In this case, the method consists at first of demonstrating nucleic acids specific to all the members of the *M. tuberculosis* complex by detection according to step (iv) of the hybridization complexes formed with a first specific nucleotide probe of the sequence SEQ ID No. 1 or of its complementary strand, then in finding among the amplified nucleic acids capable of forming complexes with the said first probe those which can likewise form complexes with a second specific nucleotide probe of SEQ ID No. 2 or of its complementary strand.

The amplified nucleic acid sequences hybridizing uniquely with the first probe correspond to specific nucleic acid sequences of the BCG, although the sequences hybridizing with each of the two probes correspond to sequences of the other members of the *M. tuberculosis* complex.

This method of differential diagnosis is of obvious interest with respect to the conventional methods of detection, because it allows an infection by BCG (possibly from a vaccination) to be differentiated from that by a virulent mycobacterium of the *M. tuberculosis* complex (*M. tuberculosis*, *M. bovis* or *M. africanum*, and possibly *M. microti*). This

10086206.0228007

- 13 -

distinction is particularly important in immunodeficient individuals, especially subjects infected by HIV.

According to another preferred embodiment, the invention provides a method of identification of groups of mycobacteria belonging to the *M. tuberculosis* complex, characterized in that:

- the DNA of the said strains previously extracted with a pair of primers such as defined above is contacted under conditions allowing a specific hybridization of the said primers with their corresponding sequences on the DNA of the said strains and the obtainment of amplification products, and

- the length of the amplification products obtained is measured, for example by agarose gel electrophoresis.

Advantageously, the primers 5'GCGCGAGAGCCCG-AACTGC3' and 5'GCGCAGCAGAAACGTCACC3' are used in this method.

The invention likewise relates to a kit for the in vitro identification of strains of mycobacteria belonging to the *M. tuberculosis* complex in a biological sample comprising:

- a pair of primers according to the invention, as defined above;

- the reagents necessary to allow the amplification of the specific sequences of nucleic acids belonging to the *M. tuberculosis* complex with the aid of the said primers,

- possibly means for revealing the amplified fragments, preferentially a nucleotide probe of the invention.

Other characteristics and advantages of the invention are illustrated by the examples following the description as well as by the figures whose legends are indicated below.

10086205.000002

- 14 -

Figure 1: "Southern Blot" analysis of the DNA of BCG (IPP). A). The fragment of 259 base pairs comprising a part of the *regX3* gene was used as a probe to detect the gene in different restriction fragments in the chromosomal DNA of the BCG (IPP). The restriction enzymes used are indicated at the top of the figure.

B). Partial restriction map of the locus of the *senX3* and *regX3* genes of the BCG (IPP). The restriction sites are indicated with respect to the *senX3* and *regX3* genes, as well as the probe used.

Figure 2: Nucleotide sequence of the *senX3* and *regX3* genes of the BCG (IPP) and derived protein sequences. The arrows indicate palindromic sequences. The putative Shine-Dalgarno sequence is underlined in dots (SD). The predicted transmembrane sequence of *SenX3* is doubly underlined. The regions of *SenX3* conserved in other sensors are underlined and annotated: H, region containing modified histidine; N, region rich in asparagine; F, region rich in phenylalanine; G1 and G2, regions rich in glycine. The small vertical arrows indicate the residues which are predicted to be phosphorylated. The annotated sequence PgmY indicates the end of the gene encoding phosphoglycerate mutase.

Figure 3: hydrophobicity profile of *SenX3*. The positive values indicate hydrophobic regions and the negative values indicate hydrophilic regions. The arrows indicate the possible initiation codons and the two horizontal lines indicate predicted transmembrane regions. The numerals at the top of the figure indicate the numbers of amino acids.

Figure 4: comparison of the intercistronic region between the BCG (IPP) and *M. tuberculosis* (IPT). The *senX3* and *regX3* genes are indicated by the arrows. The nucleotide sequences are indicated as well as the inferred protein sequences.

10086205-022802

- 15 -

Figure 5: Southern blot analysis of DNA of *M. tuberculosis* and of BCG. The chromosomal DNA of *M. tuberculosis* (IPL) (on the right) and of BCG (IPP) (on the left), was digested by *Pst*I, subjected to electrophoresis in agarose and analyzed by hybridization with the *senX3-regX3* intergenic probe of *M. tuberculosis* (IPL).

Figure 6: analysis by electrophoresis in agarose (2.5%) of the products obtained by PCR carried out on different mycobacterial strains. Lanes 1 to 3: group 1, respectively *M. microti*, *M. tuberculosis* V808, *M. tuberculosis* V761; lanes 4 and 5, group 2, respectively *M. tuberculosis* V729, *M. bovis* 60; lanes 6 to 8, group 3, respectively *M. bovis* 63, *M. bovis* 78 and *M. bovis* AN5; lanes 9 and 10, group 4, respectively *M. tuberculosis* H37Ra (IPL) and *M. tuberculosis* H37Rv (IPP); lanes 11 and 12, group 5, respectively *M. tuberculosis* (IPL), *M. bovis* 76; lanes 13 and 14, group 6, respectively *M. bovis* (BCGite 29) and *M. bovis* BCG (IPP), lane 15, group 7, *M. tuberculosis* No. 19.

Figure 7: Southern blot analysis of the fragments obtained in Figure 6. The probe used is the *senX3-regX3* inter-cistronic region of *M. tuberculosis* (IPL).

## EXAMPLES

**EXAMPLE 1:** Genomic map and cloning of the *senX3-regX3* genes of *M. bovis* BCG (IPP), coding for the two-component mycobacterial system.

Wren et al. (1992) amplified a fragment of 259 base pairs from the gene of *M. tuberculosis* (IPL) which they called *rcgX3*.

The authors of the present invention amplified the corresponding sequence from *M. bovis* BCG (vaccine strain 1173P2, obtained at the WHO collection of Stockholm, Sweden) using the following synthetic oligonucleotides: 5'-CGAGGAGTCCCTCGCCGATCCGC-3' and 5'-

- 16 -

AGCGCCCCAGCTCCAGCCGACC-3'. The amplification was carried out by using these primers and a Deep Vent polymerase (New England Biolabs). The amplification product of 259 base pairs resulting was then purified by electrophoresis on agarose gel and then sub-cloned in the *Sma*I site of the pBluescript KS+ vector (Stratagene) according to standard protocols (Sambrook et al., 1989). This insert was then removed from the recombinant plasmid by digestion with *Bam*HI and *Eco*RI and then labelled with dCTP[ $\alpha$ - $^{32}$ P] by "random priming" (random labelling) using the "random priming" kit marketed by Boehringer according to the conditions recommended by the manufacturer. The genomic DNA of *M. bovis* BCG (IPP), cultured in Sauton medium (Sauton, 1912) at 37°C in flasks for stationary culture of tissues, was extracted as described previously (Kremer et al., 1995a) and digested by different restriction enzymes. This DNA was then subjected to electrophoresis on agarose gel and to Southern Blot analysis according to the standard methods (Sambrook et al., 1989). The testing of the blot with the [ $^{32}$ P]-labelled DNA showed that digestion by *Kpn*I with either *Bam*HI, *Eco*RI or *Pst*I produced pairs of hybridizing bands, with one band labelled more intensely than the other in each of the pairs (Fig. 1a). These bands were attributed to the 5' and 3' sequences respectively, of the unique asymmetric *Kpn*I site of the probe. This allowed the *Bam*HI, *Eco*RI and *Pst*I 5' and 3' sites of the *Kpn*I site to be localized (Fig. 1b). The digestions with only *Pst*I gave a single hybridization band. As could be predicted, the digestion with *Kpn*I gave two bands and the digestion with *Bam*HI and *Eco*RI led to the obtainment of a single band of approximately 3.5 kb.

The genomic DNA of the BCG was digested with *Bam*HI and *Eco*RI, and DNA fragments of 3 to 4 kb were isolated after electrophoresis on agarose gel. These fragments were inserted into pBluescript SK (Stratagene) restricted by *Bam*HI and *Eco*RI. 900

10086206-022802



- 17 -

recombinant clones were screened with the [<sup>32</sup>P]-labelled probe by the said standard technique of "colony blot hybridization" (Sambrook et al., 1989). Three of them were positive and turn out to contain the same

5 BamHI/EcoRI DNA insert of 3.2 kb by restriction analysis. A hybridizing clone was isolated for the latter studies and was named pRegX3Bcl (I-1765).

**EXAMPLE 2:** Sequence of the genes of *M. bovis* BCG (IPP) coding for the two-component system

10 the SalI fragments of 1.0, 1.5 and 0.7 kb were isolated from the 3.2 kb insert of pRegX3Bcl (I-1765), and sub-cloned in pBluescript SK-. The sequence of the 3.2 kb insert, as well as that of the sub-clones SalI was determined on the two strands by the method of

15 "primer walking". The nucleotide sequence of 3208 bp of *M. bovis* BCG (IPP) (Fig. 2) showed the frequency of C + G characteristic of the mycobacteria (66.6%). Two principal open reading frames (ORF) designated by regX3 and senX3 were identified (Fig. 1b and 2). The regX3

20 gene starts with an ATG triplet at the 1679 position and ends with a TAG triplet at the 2360 position. It contains the sequence of the <sup>32</sup>P-labelled probe and codes for a protein whose inferred sequence of amino acids is 227 residues for a calculated molecular mass

25 of 24,881 Da. The limit upstream of the senX3 ORF is not certain because five in-phase potential initiation codons (ATG or GTG) were found in a short distance (from 296 to 446) (Fig. 2 and 3). However, only the CTC at position 296 is preceded with 9 nucleotides

30 upstream, by a sequence homologous to the Shine-Dalgarno sequence of *Escherichia coli* (4 of the 6 residues are identical to AGGAGG). Thus, this GTG codon could be the limit in 5' of the senX3 ORF which ends with the TGA stop codon at the position 1526. This ORF

35 is thus presumed to code for a protein of 44,769 Da composed of 410 amino acid residues. The senX3 gene is preceded at 273 nucleotides upstream by the 3' part of an ORF encoding a homologue of the phosphoglycerate

- 10

mutases. A sequence capable of forming a hairpin structure is localized between the positions 178 and 194. This sequence could function as a terminatory transcription sequence of the ORF upstream.

5           A search on the MycDB database showed that the *senX3* and *regX3* genes are also present in *M. leprae*. The products which they encode are respectively similar to 82.7% and 94.9% to the proteins *senX3* and *RegX3* of *M. bovis* (IPP). The *senX3* ORF of *M. leprae* turns out to  
10 be initiated by a GTC codon and is preceded by an ORF coding for a homologue of phosphoglycerate mutase, similar to that which was found in *M. bovis* BCG (IPP).

EXAMPLE 3: Cloning and sequencing of the *senX3-regX3* genes of *M. tuberculosis* (IPL).

15           The *senX3* and *regX3* genes of *M. tuberculosis* (IPL) were cloned by PCR from chromosomal DNA of *M. tuberculosis* 22962067, a clinical isolate from the collection of mycobacteria of the Institut Pasteur de Lille. The DNA of *M. tuberculosis* (IPL) was extracted  
20 by the method described above for the extraction of chromosomal DNA of the BCG (Kremer et al., 1995a). The fragment of 2.2 kb containing the *senX3* and *regX3* genes of *M. tuberculosis* (IPL) was amplified by PCR using the following primers, homologous to the adjacent sequences  
25 of the *senX3* and *regX3* genes of *M. bovis* BCG (IPP): 5'-TGGCGTAGTGTTGACTTGTC-3' and 5' GACCAGACAGTCGCCAAGGTT-3'. The amplified fragment was cloned in the *Sma*I site of pBluescript SK- (version II) to produce a plasmid named pRegX3Mt1 (I-1766). The total fragment of 2.2 kb  
30 was then sequenced using the same strategy as that described in Example 2 for the *senX3* and *regX3* genes of the BCG (IPP). The DNA sequence of the *senX3* and *regX3* ORF of *M. tuberculosis* (IPL) as well as the 5' region upstream of *senX3* and the 3' region downstream of *regX3*  
35 were identical to those of the BCG (IPP). However, the intercistronic region between *senX3* and *regX3* demonstrated interesting differences which are studied in the example below.

10085206 . 022502

- 19 -

EXAMPLE 4: Analysis of the *senX3-regX3* intercistronic region

The intercistronic region between *senX3* and *regX3* contains a perfect duplication of 77 base pairs in tandem in the BCG (IPP) (Fig. 4). Each repeated sequence contains a short ORF which has the capacity to code for a peptide of 25 amino acids. The ATG initiation codon which can be inferred from the first repetition overlaps the TGA stop codon of *senX3*. This repeated sequence ends with two stop codons in-phase, which overlap the ATG initiation codon inferred from the following repetition. The ORF of the second repetition also ends with a double TGA stop codon, which overlaps the ATG start codon of *regX3*.

The intercistronic region of *M. tuberculosis* (IPL) is longer and contains a third repeated sequence which, however, is not complete. It contains a short in-phase internal deletion of the nucleotides 40 to 66 which are replaced by GAG. The ORF of this third repetition also ends with a double stop codon overlapping the ATG codon of *regX3*.

The sequence of the intercistronic region of *M. leprae* is shorter than that of BCG (IPP). It contains 52 base pairs.

The existence of these structures in repetition in the *senX3-regX3* intercistronic region led the inventors to find out whether they were present in other regions of *M. tuberculosis* (IPL) or in the chromosome of *M. bovis* BCG (IPP) by Southern Blot analysis. The *senX3-regX3* intercistronic region of *M. tuberculosis* (IPL) was obtained by enzymatic digestion of pRegX3Mt2 with *EcoRI* and *BamHI*. The resulting DNA fragment of 491 base pairs was then digested with *BsrI*-*AluI* to produce a DNA fragment of 218 base pairs which corresponds to the *senX3-regX3* intergenic fragment of *M. tuberculosis* (IPL). This DNA fragment was then labelled at random with digoxigenin-dUTP (kit cat. No. 1093657) according to the recommendations of the

208270 90288007

- 20 -

manufacturer (Boehringer Mannheim). pReqX3Mt2 was produced by a first amplification by PCR of a fragment of chromosomal DNA of 471 base pairs of *M. tuberculosis* 22962067 (IPL) using the following oligonucleotides:  
5 5'AAACACGTCGCGGCTAATCA 3' and 5'CCTCAAAGCCCCTCCTTGCGC 3' and the resulting amplified fragment was then cloned in the SmaI site of pBluescript KS-.

The chromosomal DNA of *M. tuberculosis* (IPL) was then totally digested with PstI and subjected to  
10 electrophoresis on agarose gel and analysed by Southern Blot according to the customary procedures (Sambrook et al., 1989). The labelled probe was then used for hybridization as indicated below.

The probe was at first denatured by boiling for  
15 5 minutes, and the membrane was incubated with the probe overnight at 68°C, after a prehybridization for two hours at 68°C. The buffer used for the prehybridization and the hybridization was 5×SSC (Sambrook et al., 1989), 0.1% N-laurylsarcosine (w/v);  
20 0.02% SDS (w/v); 1% blocking reagent (Boehringer Mannheim). The membrane was then washed twice for five minutes in 2×SSC (Sambrook et al., 1989) and 0.1% SDS at ambient temperature and twice for 15 minutes in 0.1×SSC, 0.1% SDS at 68°C. The hybridized probes were  
25 then immunologically detected with antidigoxigenin Fab fragments conjugated to alkaline phosphatase and a chemiluminescent CSPD substrate (Boehringer Mannheim).

As can be seen in Figure 5, different copies of the *senX3-regX3* intercistronic region are present in *M. tuberculosis* (IPL), as well as in *M. bovis* BCG (IPP).  
30

**EXAMPLE 5:** Amplification by PCR of the *senX3-regX3* intercistronic region of different strains of mycobacteria.

Given that the intercistronic region separating  
35 the *senX3* and *regX3* genes has variations in length between *M. bovis* BCG (IPP), *M. tuberculosis* (IPL) and *M. leprae*, the authors of the present invention analysed the corresponding region in other strains of

10086206-02802

- 21 -

*M. bovis* (including non-BCG strains) and of *M. tuberculosis*. They likewise analysed other species of mycobacteria: (i) the other members of the *M. tuberculosis* complex: *M. africanum* and *M. microti* and  
5 (ii) the mycobacteria other than those of the *M. tuberculosis* complex.

The strains analysed are indicated in Tables 1 and 2. Their chromosomal DNA was prepared as indicated below. The mycobacteria were recovered by  
10 centrifugation (300 revs/min; 30 minutes), starting from 100 ml of culture and incubation for 1 hour at 37°C in 10 ml of phosphate buffer (0.4 M sucrose, 10 mM EDTA, Tris-HCl (pH 8) 10 mM, 4 mg/ml of lysozyme). The protoplasts obtained were recovered by centrifugation  
15 (3000 revs/min; 20 minutes) and then lysed by placing them to incubate for 1 hour at 60°C in 6 ml of L buffer (10 mM NaCl, SDS 6%, 10 mM Tris-HCl (pH 8), 500 µg/ml of proteinase K). After addition of 1.5 ml of 5 M NaCl, the mixture was centrifuged (14,000 revs/min., 20  
20 minutes). The supernatant was subjected to extraction with phenol-chloroform and the DNA was precipitated with isopropanol. The resuspended pellet was then treated with RNase, extracted with phenol-chloroform and chloroform and then precipitated with ethanol. The  
25 pellet was then dried in air and resuspended in 100 µl of TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 7.6).

The primers for PCR analysis were chosen from the DNA sequence of the *senX3* and *regX3* genes of *M. bovis* BCG (IPP) and *M. tuberculosis* (IPL). A primer  
30 (C5) hybridized to the 3' end of the *senX3* gene and had the following 5'-3' sequence: 5'GCGCGAGAGCCCCGAAGTGC3'. The other primer (C3) hybridized to the 5' end of the *regX3* gene and had the following 5'-3' sequence: 5'-GCGCAGCAGAAACGTCAGC-3'. With these two primers, the PCR  
35 product comprises 56 additional base pairs at the 5' end of the intergenic region and 62 additional base pairs at its 3' end with respect to the intercistronic region. The PCR product has a length of 369 base pairs

10086206.022802

- 22 -

for *M. tuberculosis* 22962067 (IPL) and 276 base pairs for *M. bovis* BCG (IPP).

The PCR amplifications were carried out in a thermocycler (Perkin Elmer) by incubation of 1 µl of chromosomal DNA with the following reaction mixture: C5 and C3 oligonucleotides (1 µg/1 µl of each), dTNP (2 µl) 25 mM, TMACl (tetramethylammonium chloride) (2 µl) 5mM, 10 x enzymatic buffer (10 µl), Vent DNA polymerase (1 U/0.5 µl), water (82 µl). The amplification was carried out over 30 cycles at 94°C for one minute, 65°C for one minute and 72°C for one minute. 10 µl of PCR product were then subjected to electrophoresis on 2.5% agarose gel and visualized with ethidium bromide. The negative controls containing all the PCR reagents except the DNA matrix were treated in parallel with the samples. The length of the PCR product was estimated by comparison with the DNA scale of 1 kb.

No PCR product was detected for the 11 strains of non-tuberculosis mycobacteria tested (see Table 2). Three strains of *Streptomyces* (*S. cacaoi*, *S. R61* and *S. R39*), as well as *E. coli* TGI likewise gave negative results in PCR. On the other hand, PCR products were obtained with all the members of the *M. tuberculosis* complex. The specificity of the PCR products was confirmed in all the cases by Southern Blot analysis using as probe the *senX3-regX3* intergenic region of *M. tuberculosis* (IPT) labelled with digoxigenin, as described in Example 4.

The length of the amplified fragments is indicated in Table 3. Eight different groups with different amplicon lengths were obtained. Out of the 35 clinical isolates of *M. tuberculosis* tested, 34 gave PCR fragments of 326 base pairs which could not be distinguished from that obtained with the reference strain *M. tuberculosis* 22962067 (IPL) (group 5). One strain of *M. tuberculosis* (No. 19) gave a PCR product

200806206-0228007

- 23 -

of 254 base pairs (group 8). *M. tuberculosis* S200 also belongs to group 5.

5 Strains of *M. tuberculosis* from Vietnam and not containing the IS6110 sequence (V. 808, V. 761 and V. 729) were different from the major *M. tuberculosis* group (group 5). The length of their PCR product exceeded 500 base pairs (one product of +/- 500 base pairs and two of 560). The laboratory strains of *M. tuberculosis* H37Ra and H37Rv had PCR products which  
10 were slightly larger than those obtained in group 5 of the strains of *M. tuberculosis* and these two strains are classified among group 4. These two strains are odd in as far as the uncertain pathogenicity for H37Rv and the loss of pathogenicity of H37Ra are concerned. In  
15 addition, no clinical case of *M. tuberculosis* was present in groups 4, 6, 8; these three groups are thus specific for the non-virulent strains.

The BCG strains (vaccine strains and isolates of clinical cases of BCGites) were divided into three  
20 groups (No. 4, 6 and 8). The PCR products obtained from these strains had a length of 353 base pairs, 276 and 199 respectively.

The most important variability at the level of the length of the PCR fragments was encountered for the  
25 non-BCG strains of *M. bovis*. Six strains were tested. Three of them, including the reference strain AN5, gave PCR products of 406 base pairs (group 3). The three other strains were grouped in groups 2, 5 and 7 corresponding to PCR fragments of approximately 500 and  
30 of 329 or 254 base pairs respectively.

10086206-022802

Species	Strains	Sources
<u>M. tuberculosis</u>	<ul style="list-style-type: none"> <li>- reference (22962067)</li> <li>- H37Ra</li> <li>- H37Rv</li> <li>- S200</li> <li>- 35 clinical isolates (different profiles of IS6110 RFLP)</li> <li>3 clinical isolates from Vietnam : V808, V761, V729</li> </ul>	<ul style="list-style-type: none"> <li>- IPL</li> <li>- IPP</li> <li>- IPP (Marchal)</li> <li>- CHR of Lille</li> <li>- CHR of Lille</li> <li>- IPP (Marchal)</li> </ul>
<u>M. bovis</u>	<ul style="list-style-type: none"> <li>- <u>BCG (vaccine strains):</u> <ul style="list-style-type: none"> <li>BCC IPP</li> <li>BCG Moreau</li> <li>BCG Japonicus</li> <li>BCG Praques</li> <li>BCG Montreal</li> <li>BCG Russe</li> <li>BCG Glaxo</li> </ul> </li> <li>- <u>case of BCGites:</u> <ul style="list-style-type: none"> <li>4 clinical isolates (28, 29, 30, 31)</li> </ul> </li> <li>- <u>non BCG:</u> <ul style="list-style-type: none"> <li>reference ANS</li> <li>1 clinical isolate (1)</li> <li>2 isolates from goats (60, 63)</li> <li>2 isolates from cows (76, 78)</li> </ul> </li> </ul>	<ul style="list-style-type: none"> <li>- IPP</li> <li>- IPL</li> <li>- IPL</li> <li>- IPP</li> <li>- IPP</li> <li>- IPP</li> <li>- IPP</li> <li>- IPL</li> <li>- CHR of Lille</li> <li>- IPP (Marchal)</li> <li>- CHR of Lille</li> <li>- University of Zaragoza (Martin)</li> <li>- University of Zaragoza (Martin)</li> </ul>
<u>M. africanum</u>	<ul style="list-style-type: none"> <li>- clinical isolate</li> </ul>	
<u>M. microti</u>	<ul style="list-style-type: none"> <li>- reference ATCC 19422</li> </ul>	<ul style="list-style-type: none"> <li>- CHR of Lille</li> <li>- IPP (Marchal)</li> </ul>

g Institut Pasteur de Lille  
" Institut Pasteur de Paris  
--- Centre Hospitalier Régional de Lille



Table 2: Atypical mycobacteria

Species and strains	Sources
<i>M. aurum</i>	- clinical isolates (IPL)
<i>M. avium</i>	- clinical isolates (IPL)
<i>M. chelonae</i>	- clinical isolates (IPL)
<i>M. flavescens</i>	- clinical isolates (IPL)
<i>M. fortuitum</i>	- clinical isolates (IPL)
<i>M. kansasii</i>	- clinical isolates (CHR of Lille)
<i>M. marinum</i>	- clinical isolates (IPL)
<i>M. scrofulaceum</i>	- clinical isolates (IPL)
<i>M. smegmatis</i>	- IPL
<i>M. terae</i>	- clinical isolates (IPL)
<i>M. xenopi</i>	- clinical isolates (IPL)

208220 90299007

Table 3: Amplification by PCR of the *senX3-regX3* intergenic region and sequencing

	Strains of mycobacteria	Estimated length of the product obtained during PCR	Composition of the <i>senX3-regX3</i> intergenic region
1	- <i>M. microti</i> - <i>M. tuberculosis</i> - V.808 - V.761	<u>560 bp</u>	77 77 77 77 77 53 -->-->-->-->-->-->-->-->-->-->
2	- <i>M. tuberculosis</i> V.729 - <i>M. bovis</i> 50 (goat)	$\pm$ 500 bp	
3	- <i>M. bovis</i> - 63 (goat) - 78 (cow) - AN5	<u>406 bp</u>	77 77 77 53 -->-->-->-->-->-->-->-->-->-->
4	- <i>M. tuberculosis</i> H37Rv (IPP) - <i>M. tuberculosis</i> H37Ra (IPL) - BCG: - <u>BCG Japonicus</u> - BCGite 28 - BCGite 30	<u>353 bp</u>	77 77 77 -->-->-->-->-->-->-->-->-->-->

2082220" 9029800F

	Strains of mycobacteria	Estimated length of the product obtained during PCR	Composition of the senX3-regX3 intergenic region
5	<ul style="list-style-type: none"> <li>- <u>M. africanum</u></li> <li>- <u>M. bovis 76</u> (cow)</li> <li>- <u>M. tuberculosis</u>:               <ul style="list-style-type: none"> <li>- reference IPL</li> <li>- S200</li> <li>- 34 clinical isolates.</li> </ul> </li> </ul>	<u>329 bp</u>	77 77 53 -->-->-->
6	<ul style="list-style-type: none"> <li>- BCG:               <ul style="list-style-type: none"> <li>- <u>IPP reference</u></li> <li>- <u>BCG Moreau</u></li> <li>- <u>BCG Russe</u></li> <li>- <u>BCG Glaxo</u></li> <li>- <u>BCGite 29</u></li> <li>- <u>BCGite 31</u></li> </ul> </li> </ul>	<u>276 bp</u>	77 77 -->-->-->
7	<ul style="list-style-type: none"> <li>- <u>M. bovis I</u></li> <li>- <u>M. tuberculosis No.19</u>: one of the 35 clinical isolates of the CHR of Lille tested</li> </ul>	<u>254 bp</u>	77 53 -->-->-->
8	<ul style="list-style-type: none"> <li>- BCG:               <ul style="list-style-type: none"> <li>- <u>BCG Praques</u></li> <li>- <u>BCG Montreal</u></li> </ul> </li> </ul>	<u>193 bp</u>	77 -->-->-->

28 -

The amplified DNA fragments comprise the *senX3-regX3* intercistronic region as well as 56 base pairs upstream of this and 62 base pairs downstream. The arrows 77 and 53 designate the repeated elements found in each group. The strains for which the intercistronic region was sequenced are underlined.

10086206 022802

BIBLIOGRAPHY

- Barnes, P.F., A.B. Bloch, P.T. Davidson, and D.E. Snider. 1991. Tuberculosis in patients with human immunodeficiency virus infection. *N. Engl. J. Med.* 324: 1644-1650.
- 10 Brisson-Noël, A., D. Lecossier, X. Nassif, B. Giquel, V. Levy-Frebaull, and A.J. Hance. 1989. Rapid diagnosis of tuberculosis by amplification of mycobacteria DNA in clinical samples. *Lancet* ii: 1069-1072.
- 15 Brisson-Noël, A., C. Aznar, C. Chureau, S. Nguyen, C. Pierre, M. Bartoli, R. Bonete, G. Pialoux, B. Cicquel, and G. Garrigue. 1991. Diagnosis of tuberculosis by DNA amplification in clinical practice evaluation. *Lancet* 338: 364-366.
- 20 Cave, M.D., K.D. Eisenach, P.F. McDermott, J.I.I. Bates, and J.T. Crawford. 1991. IS6110: conservation of sequence in the *Mycobacterium tuberculosis* complex and its utilization in DNA fingerprinting. *Mol. Cell. Probes* 5: 73-80.
- 25 Clarridge, J.E., R.M. Shawar, T. Shinnick, and B.B. Plikaytis. 1993. Large-scale use of polymerase chain reaction for detection of *Mycobacteria tuberculosis* in a routine mycobacteriology laboratory. *J. Clin. Microbiol.* 31: 2049-2056.
- 30 Cousins, D.V., S.D. Wilton, B.R. Francis and B.L. Beth. 1992. Use of PCR for rapid diagnosis of tuberculosis. *J. Clin. Microbiol.* 30: 255-258.
- 35 Eisenbach, K.D., D. Cave, J.I.I. Bates, and J.T. Crawford. 1990. Polymerase chain reaction amplification

10086206-0229007

- 30 -

- of a repetitive DNA sequence specific for *Mycobacterium tuberculosis*. *J. Infect. Dis.* 161: 977-981.
- Folgueira, L., R. Delgado, E. Palenque, and A.R. Noriega. 1993. Detection of *Mycobacterium tuberculosis* DNA in clinical samples by using a simple lysis method and PCR. *J. Clin. Microbiol.* 31: 1091-1021.
- Forbes, B.A. and K. Hicks. 1993. Direct detection of *Mycobacteria tuberculosis* in respiratory specimens in a clinical laboratory by P.C.R. *J. Clin. Microbiol.* 31: 1688-1694.
- Fries, J.W.U., R.J. Patel, W.F. Piessens, and D.F. Wirth. 1991. Detection of untreated mycobacteria by using PCR and specific DNA probes. *J. Clin. Microbiol.* 29: 1744-1747.
- Hermans, P.W.M., A.R.J. Schuilema, D. van Soolingen, C.P.H.J. Verstyne, E.M. Bik, J.E.R. Thole, A.H.J. Kolk and J.D.A. van Embden. 1990. Specific detection of *Mycobacterium tuberculosis* complex strains by PCR. *J. Clin. Microbiol.* 28: 1204-1213.
- Imaeda, T. 1985. Deoxyribonucleic acid relatedness among selected strains of *Mycobacterium tuberculosis*, *M. bovis*, *M. bovis* BCG, *M. microti*, and *M. africanum*. *Int. J. Syst. Bacteriol.* 35: 147-150.
- Jonas, J. *J. Clin. Microbiol.*, M.J. Alden, and J.I. Curry. 1993. Detection and identification of *Mycobacteria tuberculosis* directly from sputum sediments by amplification of rRNA. *J. Clin. Microbiol.* 31: 2410-2416.
- Kolk, A.H.J., A.R.J. Schuitema, S. Kuiper, V. van Leeuwen, P.W.M. Hermans, J.D.A. van Embden, and R.A. Hartskeerl. 1992. Detection of *Mycobacterium tuberculosis* in clinical samples by using PCR and a

10086206-022300

- 31 -

nonradioactive detection system. J. Clin. Microbiol.  
30: 2567-2575.

- Kox, L.F.F., J. van Leeuwen, S. Knijper, H.M. Jansen,  
5 and A.H.J. Kolk. 1995. PCR assay based on DNA coding  
for 16S rRNA for detection and identification of  
mycobacteria in clinical samples. J. Clin. Microbiol.  
33: 3225-3233.
- 10 Kremer, L., A. Baulard, J. Estaquier, O. Poulain-  
Godefroy, and C. Locht. 1995a. Green fluorescent  
protein as a new expression marker in mycobacteria.  
Mol. Microbiol. 17: 913-922.
- 15 Krerner, L., A. Baulard, J. Estaquier, J. Content, A.  
Capron, and C. Locht. 1995b. Analysis of the  
Mycobacterium tuberculosis 85A antigen promoter region.  
J. Bacteriol. 177: 642-653.
- 20 Noordhoek, G.T., A.H.J. Kolk, G. Bjune, D. Catty, J.W.  
Dale, P.E. Fine, P. Godfrey-Faussett, S-N. Cho, T.  
Shinnick, S.B. Svenson, S. Wilson, and J.D.A. van  
Emden. 1994. Sensitivity and specificity of PCR for  
detection of Mycobacterium tuberculosis: a blind  
25 comparison study among seven laboratories. J. Clin.  
Microbiol. 32: 277-284.
- Parkinson, J.S. and Kofoid, E.C. 1992. Communication  
modules in bacterial signaling proteins. Annu. Rev.  
30 Genet. 26, 71-112.
- Pierre, C., D. Lecossier, Y. Boussougnant, D. Bocart,  
V. Joly, P. Veni, and A. Hance. 1991. Use of a  
reamplification protocol improves sensitivity of  
35 detection of Mycobacterium tuberculosis in clinical  
samples by amplification of DNA. J. Clin. Microbiol.  
29: 712-717.

10086206 "02800"

- 32 -

- Saboor, S.A., N.M. Johnson, and J. McFadden. 1992. Detection of mycobacteria DNA in sarcoidis and tuberculosis with PCR. *Lancet* 339: 1012-1015.
- 5 Saiki, R.K., D.H. Gelfand, S. Stoffel, S.J. Scarf, R. Higuchi, G.T. Horn, K.B. Mullis, and H.A. Erlich. 1988. Primer-directed enzymatic amplification of DNA with a thermostable DNA polymerase. *Mol. Cell, Probes* 5: 515-219.
- 10 Sambrook J., E.F. Fritsch, and T. Maniatis. 1989. Molecular cloning: a laboratory manual, Cold Spring Harbor, N.Y.
- 15 Sauton, B. 1912. Sur la nutrition minérale du bacille tuberculeux [On the mineral nutrition of the tuberculosis bacillus]. *C.R. Acad. Sci Paris*. 155:860.
- Shankhar, S., N. Manjunath, K.K. Mohan, K. Praasad, M. Behari, G.K. Shriniwas, and K. Ahuja. 1991. Rapid diagnosis of tuberculosis meningitis by PCR. *Lancet*: 5-7.
- 20 Sjöbring, U., M. Mecklenburg, A.B. Andersen, and H. Mijörner. 1990. PCR for detection of Mycobacterium tuberculosis. *J. Bacteriol.* 169: 1080-1089.
- 25 Sudre, P., G. en Dam, and A. Kochi. 1992. Tuberculosis: a global overview of the situation today. *Bull. W.H.O.* 70: 149-159.
- 30 Telenti, A., F. Marchesi, M. Balz, F. Bally, E.C. Bottger, and T. Bodmer. 1993. Rapid identification of mycobacteria to the species level by polymerase chain reaction and restriction enzyme analysis. *J. Clin. Microbiol.* 31: 175-178.

10086206.022802



- 33 -

- van Soolingen, D., P.E.W. de Haas, P.W.M. Hermans,  
P.M.A. Groenen, and J.D.A. van Embden. 1993. Comparison  
of various repetitive DNA elements as genetic markers  
for strain differentiation and epidemiology of  
5 *Mycobacterium tuberculosis*. J. Clin. Microbiol. 31:  
1987-1995.
- van Soolingen, D., P.E.W. de Haas, J. Haagsmma, T. Eger,  
P.W.M. Hermans, V. Ritacco, A. Alito, and J.D.A. van  
Embden. 1994. Use of various genetic markers in  
10 differentiation of *Mycobacterium bovis* from animals and  
humans and for studying epidemiology of bovine  
tuberculosis. J. Clin. Microbiol. 32: 2425-2433.
- Wren, B.W., Colby, S.M., Cubberley, R.R. and Palten,  
15 M.J. 1992. Degenerate PCR primers for the amplification  
of fragments from genes encoding response regulators  
from a range of pathogenic bacteria FEMS Lett. 99, 287-  
292.

20080606 022002

- 34 -

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

## 5 (i) APPLICANT:

- (A) NAME: Institut Pasteur de Lille
- (B) STREET: 1 rue du Professeur Calmette
- (C) CITY: Lille
- (E) COUNTRY: France
- 10 (F) POSTAL CODE: 59019
- (G) TELEPHONE: 0320877900
- (H) TELEFAX: 0320877906

- 15 (A) NAME: Institut National de la Santé et  
de la recherche médicale
- (B) STREET: 101 rue de Tolbiac
- (C) CITY: Paris
- (D) COUNTRY: France
- (F) POSTAL CODE: 75013
- 20 (G) TELEPHONE: 0144236039
- (H) TELEFAX: 0145850766

(ii) TITLE OF THE INVENTION: fragments of  
specific nucleic acids of M. tuberculosis

25

(iii) NUMBER OF SEQUENCES: 2

(IV) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk
- 30 (B) COMPUTER: IBM PC compatible
- (C) OPERATING SYSTEM: PC-DOS/MS-DOS
- (D) SOFTWARE: PatentIn Release #1.0,  
Version #1.25 (EPO)

## 35 (2) INFORMATION FOR SEQ ID NO: 1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 77 base pairs

208220 90298001

- 35 -

- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) CONFIGURATION: linear

5 (ii) TYPE OF MOLECULE: DNA (genomic)

(vi) ORIGINAL SOURCE:

(A) ORGANISM: Mycobacterium tuberculosis

10 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:

ATGACCTGCG CCGACGACGA TGCAGAGCGT AGCGATGAGG TCGGGGCACC 60  
ACCCCTTGC GGGGGAGAGT GCGCGCTG 70

15 (2) INFORMATION FOR SEQ ID NO: 2:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 53 base pairs

(B) TYPE: nucleic acid

20 (C) STRANDEDNESS: single

(D) CONFIGURATION: linear

(ii) TYPE OF MOLECULE: DNA (genomic)

25 (vi) ORIGINAL SOURCE:

(A) ORGANISM: Mycobacterium tuberculosis

(xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:

30 ATGACCTGCG CCGACGACGA TGCAGAGCGT AGCGATGAGG AGGAGTGGCG  
CTG

53

10086206-02300